Sambahsa

Sambahsa (IPA: [samˈbaːsa]) or **Sambahsa-Mundialect** is an international auxiliary language (IAL) and worldlang devised by French linguist Olivier Simon.[2] It is based on the Proto-Indo-European language (PIE) and has a relatively complex grammar.[3][4] The language was first released on the Internet in July 2007; prior to that, the creator claims to have worked on it for eight years. According to a study addressing recent auxiliary languages, "Sambahsa has an extensive vocabulary and a large amount of learning and reference material".[5]

The first part of the name of the language, *Sambahsa*, is composed of two words from the language itself, *sam* and *bahsa*, which mean 'same' and 'language', respectively. *Mundialect*, on the other hand, is a fusion of *mundial* 'worldwide' and *dialect* 'dialect'.

Sambahsa tries to preserve the original spellings of words as much as possible and this makes its orthography complex, though still kept regular.[6] There are four grammatical cases: nominative, accusative, dative and genitive.[7]

Though based on PIE, Sambahsa borrows a good proportion of its vocabulary from other language families, such as Arabic, Chinese, Indonesian, Swahili and Turkish.[2]

Sambahsa	
Sambahsa-Mundialect	
Created by	Olivier Simon
Date	2007
Purpose	Constructed language • **Sambahsa**
Language codes	
ISO 639-3	(sph code proposal was rejected in 2018[1])
Glottolog	*None*
IETF	art-x-sambahsa (https://www.krea tivekorp.com/clc r/)

Phonology

Sambahsa's phonology has little to do with Proto-Indo-European phonology, though the majority of its vocabulary comes from PIE.[8] The changes from PIE are not regular, since the creator of Sambahsa has tried to avoid homophones, which would have become common after the elimination of some PIE sounds like laryngeals or some aspirated consonants. Unlike some auxlangs like Esperanto, the orthography of Sambahsa is complex yet regular and consists only of the 26 letters of the basic Latin alphabet.[9] This system was chosen to preserve the recognizability of words taken from West-

European languages, where orthography plays a key role. For example, according to the rules of Sambahsa, *bureau* is pronounced as in French, and *point* as in English.

Sambahsa's phonemic inventory has 22 consonants and nine vowels (excluding the lengthened forms of these vowels).[10] To help language learners, and because IPA symbols cannot be written with all keyboards, a special simpler system has been developed, called *Sambahsa Phonetic Transcription*, or *SPT*.

Compared to other conlangs, Sambahsa words are short, often as short as English words, and highly consonantic.[11] This latter point is in accordance with the PIE background of Sambahsa, where roots have often a consonant-vocal-consonant structure.[12]

Consonants

Consonant phonemes of Sambahsa

		Bilabial	Dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Glottal
Nasal		m		n					
Stop	voiceless	p		t			k		
	voiced	b		d			g		
Affricate	voiceless			t͡s	t͡ʃ				
	voiced			d͡z	d͡ʒ				
Fricative	voiceless	f	θ	s	ʃ	ç	x		h
	voiced	v		z	ʒ			ʁ	
Approximant		w		l		j	(w)		
Trill				r					

Vowels

Vowel phonemes of Sambahsa

	Front	Central	Back
Close	i(ː) y(ː)		u(ː)
Mid	e(ː) ø(ː)	ə	o(ː)
Open	ɛ(ː)	a(ː)	

Stress

Sambahsa's stress rules are complex but regular, and tend to follow what is often found in German or Italian. This predictability implies that all words with the same orthography are pronounced and stressed the same way as each other. Thus, for example, while German *Präsident*

and Italian *presidente* are stressed on the "ent" syllable, Sambahsa *president* is stressed on the "i", since *president* can also mean "they preside", and a final "ent" never bears the stress. This regularity of accentuation can be compared with English *president* and *to preside*, two words that bear the stress on different syllables, though they share the same origin.

Grammar

Declensions

In Sambahsa, declensions are only compulsory for pronouns. The declensions of these pronouns (demonstrative/interrogative and relative/personal) are mostly parallel, and often show similarities with their Proto-Indo-European ancestors. Thus, in all Sambahsa declensions, the neuter nominative and accusative are identical, as it was the case in PIE.[13] There are identical forms for the relative and interrogative pronouns,[14] as well as for the third person pronoun and the definite article (*the* in English).[15]

Sambahsa has two numbers (singular and plural; the dual number of PIE has not been preserved) and four grammatical genders: masculine, feminine, neuter, and undetermined. This last gender, which is an innovation from PIE, is used when a noun of uncertain or unknown gender is referred to, and, in the plural, for groups containing elements of different genders. The creator of Sambahsa introduced this non-PIE element to avoid the "gender" dispute found in Esperanto.

Gender is attributed in Sambahsa according to the "true nature" of the noun referred to, as English speakers do with *he*, *she* and *it*.

Sambahsa has four grammatical cases: nominative, accusative, dative and genitive; however, their attribution tries to be as logical as possible, and not arbitrary as in many modern Indo-European languages. The nominative is the case of the subject, and the form under which words are given in dictionaries. Except for verbs describing a movement or a position (where the appropriate prepositions ought to be used), all transitive verbs must introduce the accusative case in the first place, before an eventual dative case. However, the dependent clause of indirect speech is considered as a direct object, leading to verbs introducing an indirect object, even if there is no visible direct object.

Compare:

- *Is mi antwehrdt od is ne gwehmsiet cras*, 'He answers (to) me that he won't come tomorrow'

- *Is ne mi hat antwohrden*, 'He hasn't answered (to) me'

In Sambahsa, all prepositions trigger the accusative.[16]

The [genitive](#) indicates possession, and is used after adjectives that can introduce a [dependent clause](#).

Compare:

- *Som yakin od is ghehdsiet kwehre to*, 'I'm sure that he'll be able to do that'

- *Som yakin eysen* (genitive plural) *imkans*, 'I'm sure of his abilities'

For substantives and adjectives, there are declined "free endings" (i.e. non-compulsory) used most often in literary context for [euphonics](#) or [poetry](#). This system is inspired by the euphonic endings (*ʾiʿrāb*) found in [Modern Standard Arabic](#).

Conjugation

In Sambahsa, all [verbs](#) are regular, except *ses* ('to be'), *habe* ('to have'), and *woide* ('to know', in the meaning of French *savoir* or German *wissen*). Sambahsa verbs are indicated in dictionaries not under their [infinitive](#) form, but their bare [stem](#), because the whole conjugation can be deduced from the form of this stem. The main [tenses](#) of Sambahsa are [present](#) and [past](#), but many other tenses can be obtained through the use of affixes or [auxiliary verbs](#). Sambahsa uses the following endings, which are close to those found in many [Indo-European languages](#).[17]

Person	Present and other tenses	Past tense only
First-person singular	-o, -m (*if the verb ends with a stressed vocalic sound*) or nothing (*if the last vowel of the verb is unstressed*)	-im
Second-person singular	-s	-(i)st(a)
Third-person singular	-t	-it
First-person plural	-m(o)s	-am
Second-person plural	-t(e)	-at
Third-person plural	-e(nt) ("-nt" *is compulsory if the verb ends with a stressed vocalic sound*)	-(ee)r

Sambahsa is unusual among [auxlangs](#) because of its use of a predictable [ablaut](#) system for the past tense and passive past [participles](#). For example, *eh* within a verbal stem turns to *oh*. Other verbs that cannot use ablaut can drop their [nasal infix](#), or use an improved version of the [De Wahl's rules](#). Finally, the remaining verbs simply add the past tense endings, which are optional for verbs of the categories described above.

Therefore, this system qualifies Sambahsa as a language belonging to the [Indo-European family](#)

of languages, though it remains a constructed language.

Vocabulary

Because of its rather large vocabulary for an auxlang (as of April 2021, the full Sambahsa-English dictionary contained more than 19,500 entries[18]), it is difficult to assess the share of each language in Sambahsa's eclectic wordstock. However, the main layers are (either reconstructed or extrapolated) Indo-European vocabulary, Greco-Roman scientific and technical vocabulary (which is not discussed below, as it is more or less comparable to what is found in English) and multiple sources extending from Western Europe to Eastern Asia.

Indo-European vocabulary

The core of Sambahsa's vocabulary is undoubtedly of Indo-European origin. Only a few Sambahsa words can be traced back to pre-Indo-European times (like *kamwns*, 'chamois', cf. Basque: *ahuntz*). Many basic Sambahsa words are thus very close to their reconstructed Indo-European counterparts. See (Sambahsa/Proto-Indo-European): *eghi/*H_1*eghis* ('hedgehog'), *ghelgh/*ghelghe-* ('gland'), *pehk/*pek* ('to comb'), *skand/*skand* ('to jump'), *peungst/*pn̥k^{w}sti-* ('fist'), *wobhel/*wobhel-* ('weevil'), *gwah/*gweH₂* ('to go'), *tox/*tòksom* ('yew wood' in Sambahsa; 'yew' in PIE), *treb/*trêbs* ('dwelling'), *oit/*H_1*òitos* ('oath'), *poti/*potis* ('Sir, lord'). But less attested Indo-European vocabulary is found in Sambahsa too. For example, the common Sambahsa word for *person* is *anghen*, as in *semanghen*, 'someone, somebody', and can be derived from PIE *?*H_2*enH₁ǵh*, only found in Old Armenian *anjn* ('person') and Old Norse *angi* ('smell'). And *motic* ('hoe') may be a cognate of Old Church Slavonic *motyka* and English *mattock*.

Further development from the Indo-European background

Though Sambahsa, like any other conlang, has derivation rules, it sometimes uses backformation too. For example, the relation between Lithuanian *bendras* ('companion'), Old Greek *pentheros* ('father-in-law') and Sanskrit *bandhu-* ('companion') is uncertain;[19] however Sambahsa "reconstructs" this root as *behndwr* from *behnd* 'to bind'. PIE has *dhéǵhom* 'earth' and *dhinéǵh-* (with nasal infix) 'to shape, to make pottery'; accordingly, Sambahsa has *(di)ghom* and *dinegh*, but the latter can be understood as "to put earth on" if we refer to *yug* ('yoke') and *yuneg* ('to join'), both from PIE *yugom* and *yunég-*.

The Sambahsa word for 'ice pellet' is *kersnit*; it rests on the word *kersen* 'frozen snow', itself from Old Norse *hjarn*, Lithuanian *šarma* ('frost') and Russian *serën*.[20] But the suffix *-it* was abstracted

from PIE words like *sepit* 'grain of wheat' and *H₂elbit* 'grain of barley';[21] thus *kersnit* can be understood as 'a grain of frozen snow'.

Words common to different language families

A characteristic of Sambahsa is to include words found in different language families, while the most famous auxiliary languages tend to limit themselves to a compilation of Romance vocabulary with some borrowings from the Germanic languages. For example:

- *Schkaf* ('cupboard') has cognates both in Germanic and Slavic languages: Russian шкаф, Polish *szafa*, Ukrainian шафа, Danish *skab*, Icelandic *skápur*, Franconian dialect *schaaf* and Swedish *skåp*.

- *Graf* ('count', as a nobility title) is a German word from Greek *grapheùs* that has been borrowed into many languages including Azerbaijani *qraf*, Bulgarian граф, Czech *hrabě*, Danish *greve*, Estonian *krahv*, Croatian *grof*, Hungarian *gróf*, Finnish *kreivi*, Lithuanian *grafas*, Icelandic *greifi* and Russian граф.

- *Bicair* ('mug') is found in German *Becher* and many other Germanic languages. It comes from Low Latin *bicarium* and is at the origin of Hungarian *pohár*, Italian *bicchiere* and Romanian *pahar*, all meaning 'glass'.

- *Saray* means 'big hall, palace' and has the same Turkish and Persian origin as English *seraglio* but with a meaning closer to its etymology and to Russian сарай ('barn').

The Balkan sprachbund

Though they belong to different language families, the languages spoken in Southeast Europe share a number of common grammatical features and of loanwords due to their historical background. That is why Sambahsa includes words from this region.

- *Schut* ('hornless') corresponds to Romanian *șut*, Bulgarian/Serbo-Croatian *šut*; also Albanian *shut* 'hornless'.

- *Potire* ('pitcher') comes from Old Greek ποτήρ, like Serbo-Croatian путир, Russian потир, Romanian and Albanian *potir*.

- *Keramide* ('coating') comes from Greek κεραμίδα, which has given, among others, Romanian *cărămidă* ('brick') and Arabic قرميدة *qirmîda(t)* 'tile'.

Words from Arabic and Persian

A significant part of Sambahsa's vocabulary comes from Arabic and Persian. Both languages have extensively provided loanwords to a lexical continuum ranging from the Atlantic Ocean to

Indonesia because, respectively, of the spread of Islam and the brilliance of the former Persian civilization. Sambahsa learning materials often call this stratum "Muslim".

- *Amlak* ('assets') comes from Arabic أملاك and is found in Turkish *emlak* ('estate') and Persian املاک.

- *Zina* ('adultery') comes from Arabic زنا and is found in Persian and many other languages spoken by a majority of Muslims.

- *Adarb* ('merlon') comes from Spanish *adarve* and Portuguese *adarve* from Arabic درب and ultimately Persian در which has its origin in PIE **dʰwer-* like Sambahsa *dwer* ('door').

Sinitic vocabulary

Classical Chinese has heavily influenced the wordstock of neighbouring languages, mostly Japanese, Korean and Vietnamese. As a result, Sambahsa incorporates some Sinitic vocabulary, but the phonetic differences between these various languages can be high.

- *Kjingyow* ('goldfish') corresponds to 金魚, which is read *jīnyú* in Mandarin Pinyin and *kingyo* in Japanese.

- *Geong* ('fortified palace') corresponds to the Han character 城 read *chéng* in Mandarin Pinyin, *jō* in Japanese Go-on reading, *seong* in Korean, and *thành* in Vietnamese.

- *Rjienrlwey* ('humankind') in an attempt to balance Mandarin *rénlèi*, Japanese *jinrui*, Korean *illyu*, and other renderings of 人類. This word showcases some seeming flaws of the language's approach to be reminiscent of all targeted languages at once.

Not all Sambahsa Sinitic words come from Classical Chinese. The Min Nan language of Southern China provided loanwords to some South-East Asian languages, and some of these borrowings are, in turn, found in Sambahsa.

- Sambahsa *pangsit* ('wonton') is an Indonesian word from Min Nan *pian sit*, while Mandarin Chinese (Pinyin) has *húndùn*

- Likewise, Sambahsa *loteng* ('attic') comes from Min Nan *lauteng* through Indonesian *loteng*.

Sample phrases

Sambahsa	English
Sellamat!	Hello!
Kam leitte yu?	How are you?
Leito.	Fine.
Bahte yu Sambahsa?	Do you speak Sambahsa?
No, ne bahm Sambahsa.	No, I don't speak Sambahsa.
Marba!	Pleased to meet you!

Literary works translated into Sambahsa

- The Songs of Bilitis by Pierre Louÿs : *Ia Songvs as Bilitis* (https://fr.scribd.com/doc/83065380/Ia-Songvs-as-Bilitis)

- Demian by Hermann Hesse : *Demian* (https://fr.scribd.com/doc/122334083/Demian-Sambahsa-tarjem)

- The Stranger by Albert Camus : *Is Gospoti* (https://fr.scribd.com/doc/81520739/Is-Gospoti-l-Etranger-in-Sambahsa)

- The Little Prince by Antoine de Saint-Exupéry : *Is Lytil Prince* (https://fr.scribd.com/doc/61098743/Is-Lytil-Prince-Le-Petit-Prince-en-Sambahsa-The-Little-Prince-in-Sambahsa)

- The Gospel of Matthew : *Id Euanghelio sekwent Matyah* (https://fr.scribd.com/doc/147341646/EUANGHELIO-SEKWENT-MATYAH)

- Alice's Adventures in Wonderland : *Ia Aventures as Alice in Daumsenland* (http://www.evertype.com/books/alice-sambahsa.html) and *Through the Looking-Glass* (http://www.evertype.com/books/looking-glass-sambahsa.html) published by Evertype

- The Strange Case of Doctor Jekyll and Mister Hyde : *Id Stragno Fall om Doctor Jekyll ed Poti Hyde* (https://www.evertype.com/books/jekyll-sambahsa.html) published by Evertype

- *Un Coeur simple* by Gustave Flaubert : *Un simple kerd* (https://fr.scribd.com/document/380965322/Un-Simple-Kerd)

- The Tower of the Elephant, The Scarlet Citadel, The Devil in Iron, A Witch Shall Be Born, Jewels of Gwahlur and Shadows in Zamboula by Robert E. Howard : *Id Tor ios Elephant* (https://fr.scribd.com/document/334053099/Id-Tor-Ios-Elephant) , *Id Scarlato Citadell* (https://fr.scribd.com/document/335501303/Id-Scarlato-Citadell) , *Diabel ex Sider* (https://fr.scribd.com/document/383775991/Diabel-Ex-Sider) , *Gnahsiet un Yasa* (https://fr.scribd.com/document/580718626/Gnahsiet-Un-Yasa) , *Ia Dents os Gwahlur* (https://fr.scribd.com/document/611394629/Ia-Dents-Os-Gwahlur) , *Zamboulas Anthropophags* (https://fr.scribd.com/document/638702165/Zamboulas-Anthropophags)

- A Scandal in Bohemia by Arthur Conan Doyle : *Un Scandal in Chekhia* (https://fr.scribd.com/document/469774223/Scandal-in-Chekhia)

- The Call of Cthulhu and the Moon-Bog by H. P. Lovecraft : *Kal os Cthulhu* (https://fr.scribd.com/document/558075126/Kal-Os-Cthulhu) , *Luna-moor* (https://fr.scribd.com/document/558075759/Luna-moor)

Movies with Sambahsa subtitles

- *Revelations* (a fan-made movie based on Star Wars): *Revelations* (https://www.youtube.com/watch?v=m-qjqQwF0Bc)

- The Hunt for Gollum (a fan-made prequel to the Lord of the Rings) : *Sayd po Gollum* (https://www.youtube.com/watch?v=GkAxzTfcMbU)

- Born of Hope (a fan-made prequel to the Lord of the Rings) : *Gnaht Speh* (https://www.youtube.com/watch?v=cY4aAtSf8eI)

- *Home* (a French movie by Yann Arthus-Bertrand about environmental threats) : *Ghom* (https://www.dailymotion.com/playlist/x1ikyf_3abductee_ghom-in-sambahsa-hd/1#video=xgb2g3)

- Kaydara (a fan-made movie based on *The Matrix*) : *Kaydara* (https://www.dailymotion.com/video/xickar_kaydara-official-film_shortfilms)

References

1. "Change Request Documentation: 2017-036" (https://iso639-3.sil.org/request/2017-036) . *SIL International*.

2. Olivier Simon (2010). "The Official Website of Sambahsa" (http://sambahsa.pbworks.com/w/page/10183084/FrontPage) . Retrieved 2011-02-18.

3. Mithridates (2009-05-14). "Why You Should Keep an Eye on Sambahsa" (http://www.pagef30.com/2009/05/why-you-should-keep-your-eye-on.html) . Retrieved 2011-02-18.

4. "sambahsa / FrontPage" (http://sambahsa.pbworks.com/w/page/10183084/FrontPage) . *sambahsa.pbworks.com*. Retrieved 2019-10-31.

5. "The Representation of Korean and Other Altaic Languages in Artificial International Auxiliary Languages" in *Journal of Universal Language*, March 2012, p.153, by Alan Reed Libert.

6. A full analysis of Sambahsa (written in Esperanto) has been made by S.Auclair in *La Riverego* n°104, pp. 11-16, http://www.esperanto.qc.ca/files/riverego/Riverego-104.pdf

7. Dave MacLeod (2010). "Foreword to the Sambahsa Grammar in English" (http://sambahsa.pbworks.com/w/page/27871608/Foreword-to-the-Sambahsa-Grammar-in-English) . Retrieved 2011-02-02.

8. "sambahsa / Sambahsa pronunciation in English" (http://sambahsa.pbworks.com/w/page/27811711/Sambahsa-pronunciation-in-English) . *sambahsa.pbworks.com*. Retrieved 2021-08-21.

9. "The strange quest for a universal "Earth Standard" language" by Esther Inglis-Arkell, 08-17-2012 : http://io9.com/5935563/the-strange-quest-for-a-universal-earth-standard-language

10. However, different versions of pronunciation of "r" are admitted, and the "ng" sound (as in English "sing") could be counted as a new sound, distinct from the conjunction of [n] + [g].

11. See this link on a French-speaking forum : http://aphil.forumn.net/t844p15-analyse-phonotactique-kotava-esperanto-uropi-et-autres?highlight=analyse+phon%E9tique

12. Benveniste, Émile (1962). *Origines de la formation des noms en indo-européen* (https://books.google.com/books?id=OD4IAQAAIAAJ) (in French). Adrien-Maisonneuve.

13. R.S.P. Beekes, *Comparative Indo-European Linguistics*, J.Benjamins.Pub., p.195

14. With the exception of the nominative singular masculine, as in Latin, where the relative pronoun is *qui*, and the interrogative form is *quis*.

15. But the genitive form serves only for the definite article, while the possessive pronouns have special forms (otherwise, confusions could have arisen).

16. Under certain circumstances, the preposition *bi* can merge with the definite article in its dative form.

17. They can be compared to the data provided in *Indo-European Linguistics : an introduction* by J. Clackson, Cambridge University Press, 2007, pp. 127 & 128.

18. "Sambahsa English Dictionary - Nature" (https://www.scribd.com/doc/39063111/Sambahsa-English-Dictionary). *Scribd*. Retrieved 27 June 2017.

19. J.P Mallory & D.Q. Adams, *Encyclopedia of Indo-European Culture*, Fitzroy Dearborn Publishers, p.196

20. *ibidem*, p.287

21. *ibidem*, p.639

External links

- Official Website of Sambahsa (http://sambahsa.pbworks.com/w/page/10183084/FrontPage%7CThe)

- A Sambahsa English Dictionary (https://www.scribd.com/doc/39063111/Sambahsa-English-Dictionary)

- Sambahsa-English dictionary on Glosbe (providing dictionaries to many other languages) (https://glosbe.com/mis_sam/en/)

- Sambahsa-mundialect Wiki (https://sambahsa.fandom.com/wiki/Sambahsa-mundialect_Wiki)